Legal Knowledge and Information Systems
A. Wyner and G. Casini (Eds.)

doi:10.3233/978-1-61499-838-9-107

Balancing with Thresholds

Michał ARASZKIEWICZ [a,1], Tomasz ZUREK [b]
[a] *Departament of Legal Theory, Faculty of Law and Administration, Jagiellonian University, Ul. Gołębia 24, 31-007 Kraków, Poland*
[b] *Institute of Computer Science, Maria Curie Sklodowska University, Ul. Akademicka 9, 20-033 Lublin, Poland*

Abstract The paper presents a general formal framework representing the role of balancing of values in interpretation of statutory rules. The model developed here is an extension of the model of teleological interpretation, where a given interpretive outcome is justified if it satisfies a given goal (or a set of goals). Herein, a richer argumentative structure is discussed: an interpretive proposition concerning the interpretation of a statutory condition is justified if it is in accordance with the proper balance of applicable legally relevant values.

Keywords. Argumentation, Balancing, Goal-based reasoning, Statutory Interpretation, Value-based reasoning

1. Introduction

The topic of balancing of values and its role in legal reasoning has been the point of interest in AI and Law for more than two decades now. However, so far the fundamental concepts and inference patterns related to balancing have not been accounted for in a formal framework in the context of statutory interpretation. This paper aims to fill this gap. The paper does not deal with the structure of the balancing itself, but it argues for a basic conceptual scheme that creates the background for any instance of balancing-based interpretation. The results of our work may be useful for the development of rule-based systems involving the notion of interpretation.

In legal literature the topic of balancing has been initially associated with the notion of legal principles [7], [1], [2], [10], [3]. In the domain being the scope of this paper – that is, statutory interpretation – the issue in question is whether a given rule should be interpreted in certain manner and, as a consequence, applied to the given state of affairs. Therefore, the objects being valuated with respect to relevant values are states of affairs with attached consequences following from the rule in question, and, for comparison, the same states of affairs without such consequences (similarly to [12] and [8]). As far as the criteria of acceptance are concerned, two aspects have to be distinguished. First, no legally relevant value should be realized below its core threshold [10]. Second, it is an open question whether we are obligated to adopt the interpretation which yields the optimal level of balancing of values, or is it acceptable to adopt any outcome which satisfies

[1]Corresponding Author: Departament of Legal Theory, Faculty of Law and Administration, Jagiellonian University, Ul. Gołębia 24, 31-007 Kraków, Poland; E-mail: michal.araszkiewicz@uj.edu.pl The writing of this paper was supported by the Polish National Science Centre (research project No DEC-2015/17/B/HS5/00457)

a certain valuation threshold. Our thesis is that this threshold of minimal acceptability is typically fixed by means of interpretive propositions based on balancing. Our paper is based on the idea which can be seen as the development and discussion of the teleological interpretation concept from [12], but also as an implementation of the concept of goal and relationship between goal and value from [11].

2. The model

We will begin with a summarized discussion of the basic concepts of the model of teleological reasoning from [11], further referred to as the GVR model:
Let $S = \{s_x, s_y, s_z, ...\}$ be a finite, non-empty set of propositions. Each proposition represents one state of affairs. We have to separate the two meanings of the word value: a value may be understood as a concept or as a process: (1) Value as an abstract concept which allows for the estimation of a particular action or a state of affairs and influences one's behaviour. V is a set of values: $V = \{v_1, v_2, \dots v_n\}$ (2) Valuation as a process of estimation of the level of extent to which a particular states of affairs s promotes a value v_i. By $v_i(s)$ we denote the extent to which s promotes a value v_i. By $V(S)$ we denote the set of all valuations of all states of affairs. By $V^i(S)$ we denote the set of all possible extents to which a value v_i from set V may be promoted by any possible state of affairs $s \in S$. A partial order $O_i = (\geqslant; V^i(S))$ represents the relation between extents to which values are promoted.In real-life reasoning people do not rely only on a comparison of the levels of promotion of one value; usually, they compare the levels of promotion of various values. Theoretically speaking, they are incompatible, but practically, people compare not only the levels of promotion of various values, but also the levels of promotion of various sets of values. By $V^Z \subset V$ we denote a subset (named Z) of a set of values V which consists of values: $v_i, v_j, ... \in V^Z$. By $V^{s_i} \subset V$ we will denote a set of values promoted by a state of affairs s_i.
By $V^Z(s_n)$ we denote a set of estimations of the levels of promotion of values constituting set V^Z by a state of affairs $s_n \in X$. If $V^Z = \{v_z, v_t\}$, then $V^Z(s_n) = \{v_z(s_n), v_t(s_n)\}$. A partial order $OR = (\rhd; 2^{V(S)})$ represents a preference relation between various sets of values and various states of affairs: $V^Z(s_n) \rhd V^Y(s_m)$ means that the extent to which values from set V^Z are promoted by a state of affairs s_n is preferred to the extent to which values from set V^Y are promoted by a state of affairs s_m.
The discussion of relationships between orders OR and O as well as the mechanism of deriving order OR is presented in the [11].

Definition 1 (Legal rule) *Let $R = \{r_v, r_z, ...\}$ be a set of legal rules. Each rule is a pair $\langle s_x, c_x \rangle$, where s_x is the condition of the rule and c_x is the conclusion of the rule.*
If a state of affairs s_a fulfills the conditions of the rule, then the conclusion leads to the change of the state of affairs into s_{x+c} (where $s_x, s_{x+c} \in S$).

Basing on [10] we assume that the grounds for evaluation of each interpretation in our model will be goals in the form of minimal extents to which a given set of values should be promoted. Thus established concept of goal remains complacent with the idea of abstract goal from [11]:

Definition 2 (Goals) *Goals are represented by the minimal acceptable extents to which a particular state od affairs promotes a given set of values:*

Let $GA = \{ga_1, ga_2, ...\}$ be a set of goals. By $v_n min(ga)$ we denote the minimal extent to which the promotion of a value v_n satisfies a goal ga. By $v_n(s_1) \geqslant v_n min(ga)$ we denote that a goal ga is satisfied by a state of affairs s_1 with respect to a value v_n. By $v_n \in ga$ we denote that the minimal extent of a given value v_n is declared in a goal ga (note that $\in$ is different that $\in$). The abovementioned definition of goals represents the idea of protection of the core of values; the defined goals correspond to core thresholds as discussed in [10]. Other types of goals are also relevant in law, but the minimal thresholds are particularly important, defining the minimal acceptability of statutory interpretation statements.
Although like Sartor we assume that the foundation for setting goals are values whose promotion is recommended by principles (goal norms).

Definition 3 (Interpretation) *The binary operator $\bullet$ represents interpretation of the principle's conditions (the operator $\bullet$ was extensively discussed in [4] and [5]). By $s_t \bullet s_x$ (where $s_t, s_x \in S$) we mark that a state of affairs s_t fulfills conditions s_x.*

It should be pointed out that in our model we introduced a differentiation between the current state of affairs and the state of affairs expressed in the rule's premises. It results from the fact that in practice the description of the actual state of affairs very rarely literally matches the premises of the rule; most frequently it is somehow interpreted, often by the so-called intermediate legal concepts.

Definition 4 (Interpretive Statements) *All complex expressions of the elements of set S and constructed by means of the relation word $\bullet$ will be referred to as Interpretive Statements.*

Interpretive Statements play a role of intermediaries between the factual description of a given state of affairs and the states of affairs expressed in the conditions of legal rules. The crucial question is this context is whether the conditions of a rule should be interpreted in such a way to encompass the current fact situation, or to the contrary. Interpretive canons [9] serve as arguments for justification of this or another Interpretive Statements concerning the conditions of the rule in question. However, these canons may also be looked at as heuristics: simplified rules approximating the actually justified Interpretive Statements. If we agree that law is a system designed for the sake of realization of important social values, then we may assume that the set of "actually justified" Interpretive Statements follow from the balancing of those socially relevant values. Note that we do not claim the existence of a unique "right" interpretation of any legal rule [7].
Let us now consider application of a given rule $r_a\langle s_a, c_a\rangle$ to the state of affairs s_m. In order to justify this application, we have to be able to show that $s_m \bullet s_a$ (for the sake of simplicity we do not consider the problem of analogous application of rules here). Typically, justifying this inference step will involve at least one layer of intermediary concepts.
Let us define the set $IS\ (r_a, s_m)$ as the set of Interpretive Statements concerning apppli-cation of r_a to s_m.

Definition 5 (Positive and Negative Interpretive Arguments) *An Interpretive Statement $\in IS(r_a, s_m)$ is a Positive Interpretive Statement ($PINS$) if and only if it justifies application of r_a to s_m. An Interpretive Statement $\in IS(r_a, s_m)$ is a Negative Inter-*

pretive Statement Statement ($NENS$) if and only if it justifies non-application of r_a to s_m.

Definition 6 (Goal-admissible Interpretive Statement) *An interpretive statement $\in IS(r_a, s_m)$ is goal-admissible with respect to goal ga warranting the realization of value v_n if and only if:*

a. $v_n(s_m) \geqslant v_n min(ga)$ if the interpretive statement is a $NENS$ (goal- admissible $NENS$)
b. $v_n(s_{m+c_a}) \geqslant v_n min(ga)$ if the interpretive statement is a $PINS$ (goal - admissible $PINS$)

If an interpretive statement is not goal-admissible, then it is goal-inadmissible.

It is worthwhile to delimit the set of states of affairs in which each Interpretive Statement, concerning application of legal rules R, is goal-admissible with respect to the value v_n and the goal ga, setting its minimal required realization. We will refer to this set as the Model of Interpretive Statements with regard to value v_n and goal g_a.

Definition 7 (Model of Interpretive Statements – single value) . *Let IS be the given set of Interpretive Statements, v_n the value in question and ga – the goal protecting the minimal realization of the value. $Mod(IS)_{v_n,ga}$ is the set of all states of affairs in which all given Interpretive Statements are goal-admissible.*

Intuitively, the model of interpretive statements is the set of all states of affairs where the degree of realization of a value is always greater than the minimal threshold, taking into account the rules applicable to these states of affairs and interpretation of these rules. A given state of affairs may not belong to the model of interpretive statements if this state of affairs is untypical or novel (hard cases), or if the legislation is poorly drafted, enabling goal-inadmissible interpretations.
Conversely, we may define the set of all goal-admissible interpretive statements for a given set of states of affairs, taking into account value v_n and goal ga. We will refer to this set as the Interpretive Theory ($INTh$) of a set of states of affairs.

Definition 8 (Interpretive Theory of States of Affairs – single value) *Let S be the given set of States of Affairs, v_n the value in question and ga – the goal protecting the minimal realization of the value.*
$INTh(S)_{v_n,ga}$ is the set of all Interpretive Statements that are goal-admissible in any of the states of affairs in S.

The operator $INTh(S)_{v_n,ga}$ separates the goal-admissible from the goal-inadmissible interpretive statements, for a given set of states of affairs, taking into account the realization of v_n with regard to ga.
Let us now combine the two operators to obtain the notion of Value-based Consequence of the given set of Interpretive Statements.

Definition 9 (Value-Based Consequence) *Value-based consequence of the set of Interpretive Statements $VCn(IS)_{v_n,ga}$ is defined as $INTh_{v_n,ga}(Mod(IS)_{v_n,ga})$, that is, the set of all Interpretive Statements that are goal-admissible in all non-hard cases.*

The above definitions may be generalized to encompass sets of values and goals. Intuitively, $VCn(IS)_{v_n,ga}$ defines a relatively narrow subset of Interpretive Statements.

3. Argumentation schemes

We have already noticed ([9], [4]) that interpretive statements are justified by means of interpretive canons, usually expressed as argumentation schemes which are forms of argument which represent stereotypical patterns of human reasoning.
Below we present two interpretive canons: the first one justifies a positive interpretive statement on grounds that it fulfill the goal set by the legislator; the second one constitutes a demonstration of balancing-based interpretive conflict solution.

IAS1 The first type of argumentation scheme: every positive interpretive statement which fulfills the goal is justified. The given data are: a goal ga_k, a current state of affairs (s_m), a legal rule r_l, and a Positive Interpetive Statement: $s_m \bullet s_l$. If after the application of rule (s_{m+c}), s_m will promote all values indicated by goal ga_k to a no lesser degree than the minimum, then the interpretive statement $s_m \bullet s_l$ will be justified:

$$\frac{\begin{array}{c} ga_k \\ s_m \\ r_l = \langle s_l, c \rangle \\ s_m \bullet s_l \in PINS \\ \forall_{v_n \in ga_k} v_n(s_{m+c}) \geqslant v_n min(ga_k) \end{array}}{s_m \bullet s_l}$$

IAS2 The second argumentation scheme refers directly to the balancing of values and is an example of a conflict resolution mechanism: there are two exclusive interpretive statements, both fulfilling the set goal, but one of them is preferred because of the values it promotes: The given data are: a goal ga_k, a current state of affairs (s_m), a legal rule r_l, two interpretive statements $s_m \bullet s_l \in PINS$ and $s_m \not\bullet s_l \in NENS$. Both cases promote values indicated by ga_k to the extent no lesser than the recommended minimum. If in the context of set V^k and after the application of rule (s_{m+c}), s_m will be preferred to interpretation $s_m \not\bullet s_l$, then the interpretive statement $s_m \bullet s_l$ will be justified. The crucial point in the discussion is the list of values V^k determining on the basis of which values the balance should be made. Obviously, not all of the values ought to be taken into consideration. In previous sections we assumed that our goal (ga_k) is set on the basis of the binding legal principles; since they define the constitutional order, we believe that they should serve as the foundation of balancing. Therefore we assume that $V^k = \{v_n | v_n \in ga_k\}$.

$$\frac{\begin{array}{c} ga_k \\ s_m \\ r_l = \langle s_l, c \rangle \\ s_m \bullet s_l \in PINS \\ s_m \not\bullet s_l \in NENS \\ \forall_{v_n \in ga_k} v_n(s_{m+c}) \geqslant v_n min(ga_k) \\ \forall_{v_n \in ga_k} v_n(s_m) \geqslant v_n min(ga_k) \\ V^k(s_{m+c}) \rhd V^k(s_m) \end{array}}{s_m \bullet s_l}$$

Let us note that the argumentation schemes presented above do not have to lead to the conclusion concerning uniqueness of the best interpretive sentence (one right answer, see [9]), because set ordered by the symbol ($\rhd$) may not have the greatest element, it may have more than one maximal elements.

4. Discussion and Conclusions

In this paper we have provided a general framework which extends the concept of teleological reasoning to represent the role of balancing of values in the context of statutory interpretation. This framework is compatible with the findings of [10]. We have also defined certain specific concepts representing safe interpretive situations, where explicit balancing is not needed to justify a satisfactory interpretive outcome (the notions of model of Interpretive Statements, interpretive theory of states of affairs and finally, value-based consequence). For situations where actual balancing needs to be made explicit, we have provided two argumentation schemes.

As for the future work, we intend to: (1) explore the structure of arguments supporting conclusions encompassing the ordering operators (orders O_i and OR); this line of reserach involves investigations into case-based reasoning structures in statutory interpretation; (2) apply the framework to model the situation of justified violation of rules (as in Bench Capon [6]) and (3) integrate the model into the broader framework modeling the behaviour of agents interpreting statutes.

References

[1] R. Alexy. *A Theory of Constitutional Rights*. Oxford Press, 2003.

[2] Michal Araszkiewicz. Balancing of legal principles and constraint satisfaction. In *Legal Knowledge and Information Systems - JURIX 2010: The Twenty-Third Annual Conference on Legal Knowledge and Information Systems, Liverpool, UK, 16-17 December 2010*, pages 7–16, 2010.

[3] Michał Araszkiewicz. Argument structures in legal interpretation: Balancing and thresholds. In Thomas Bustamante and Christian Dahlman, editors, *Argument Types and Fallacies in Legal Argumentation*, pages 129–150. Springer International Publishing, 2015.

[4] Michal Araszkiewicz and Tomasz Zurek. Comprehensive framework embracing the complexity of statutory interpretation. In *Legal Knowledge and Information Systems - JURIX 2015: The Twenty-Eighth Annual Conference, Braga, Portual, December 10-11, 2015*, pages 145–148, 2015.

[5] Michal Araszkiewicz and Tomasz Zurek. Interpreting agents. In *Legal Knowledge and Information Systems - JURIX 2016: The Twenty-Ninth Annual Conference*, pages 13–22, 2016.

[6] Trevor J. M. Bench-Capon. Value-based reasoning and norms. In *ECAI 2016 - 22nd European Conference on Artificial Intelligence, 29 August-2 September 2016, The Hague, The Netherlands - Including Prestigious Applications of Artificial Intelligence (PAIS 2016)*, pages 1664–1665, 2016.

[7] R. Dworkin. *Taking Rights Seriously. New Impression with a Reply to Critics.* Duckworth, 1978.

[8] Matthias Grabmair and Kevin D. Ashley. Argumentation with value judgments an example of hypothetical reasoning. In *Proceedings of the 2010 conference on Legal Knowledge and Information Systems: JURIX 2010: The Twenty-Third Annual Conference*, pages 67–76, Amsterdam, The Netherlands, The Netherlands, 2010. IOS Press.

[9] F. Macagno, D. Walton, and G. Sartor. Argumentation schemes for statutory interpretation. In M. Araszkiewicz, M. Myška, T. Šmejkalova, J. Šavelka, and Škop M., editors, *Argumentation. International Conference on Alternative Methods o Argumentation in Law.*, pages 61–76, Brno, 2012.

[10] G. Sartor. Doing justice to rights and values: teleological reasoning and proportionality. *Artificial Intelligence and Law*, 18, 2010.

[11] Tomasz Zurek. Goals, values, and reasoning. *Expert Systems with Applications*, 71:442 – 456, 2017.

[12] Tomasz Zurek and MichałAraszkiewicz. Modeling teleological interpretation. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, ICAIL '13, pages 160–168, New York, NY, USA, 2013. ACM.